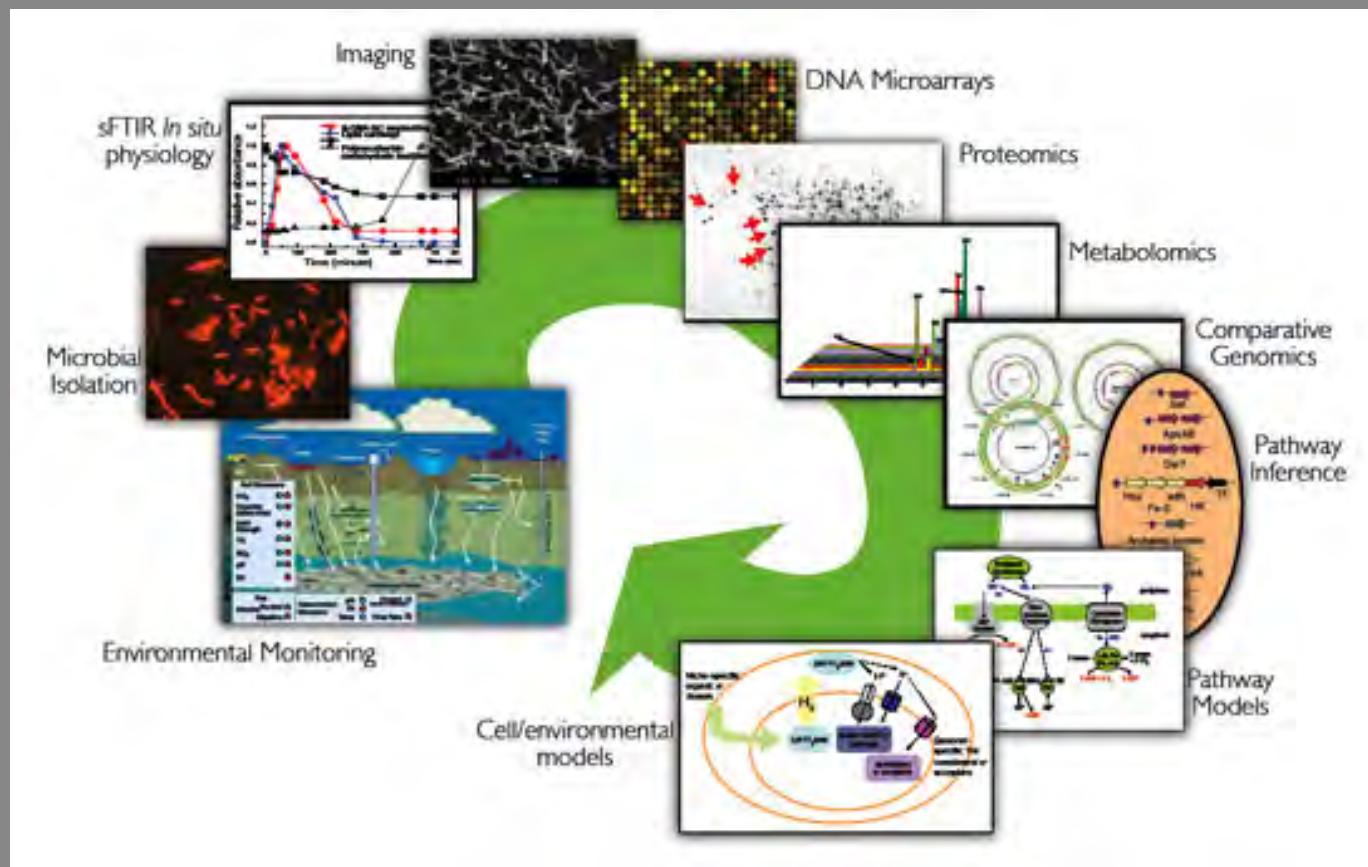


- Dashboard Design
- Team Science Approach
- Project Milestones &
- WBS & Budgets
- Working Dashboards



Amy Shutkin, CPM
ashutkin@lbl.gov
Ernest Orlando Lawrence Berkeley National
Laboratory
Physical Biosciences Division

ACKNOWLEDGEMENT

ESPP² is part of the Virtual Institute for Microbial Stress and Survival supported by the U. S. Department of Energy, Office of Science, Office of Biological and Environmental Research, Genomics Program:GTL through contract DE-AC02-05CH11231 between Lawrence Berkeley National Laboratory and the U. S. Department of Energy.

Information Dashboard Design: The Effective Visual Communication of Data

INFORMATION DASHBOARD DESIGN

The Effective Visual Communication of Data



Stephen Few

O'REILLY®

About the Author

Stephen Few has worked for over 20 years as an IT innovator, consultant, and teacher. Today, as Principal of the consultancy Perceptual Edge, Stephen focuses on data visualization for analyzing and communicating quantitative business information. He teaches in the MBA program at the University of California, Berkeley. You can learn more about Stephen's work and access an entire library of articles at www.perceptualedge.com.

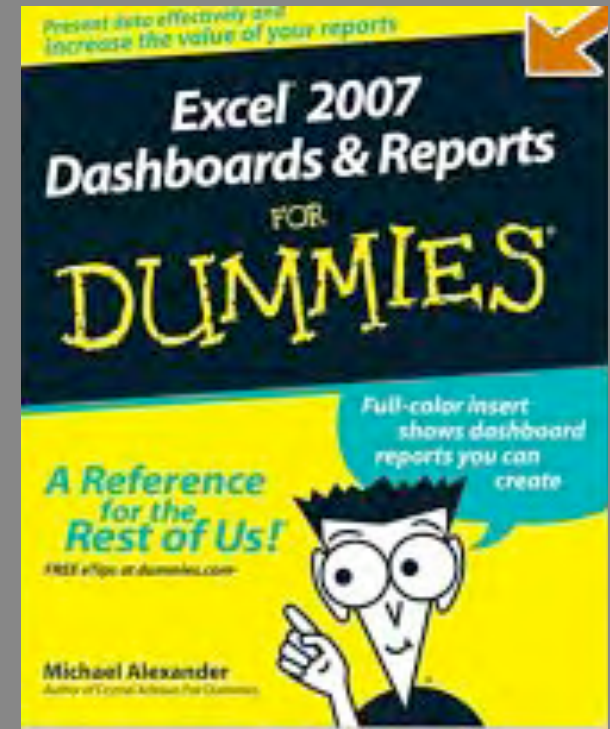
Dashboards - defined

“A dashboard is a visual interface that provides at-a-glance views into key measures relevant to a particular objective or business process.”

Key Attributes:

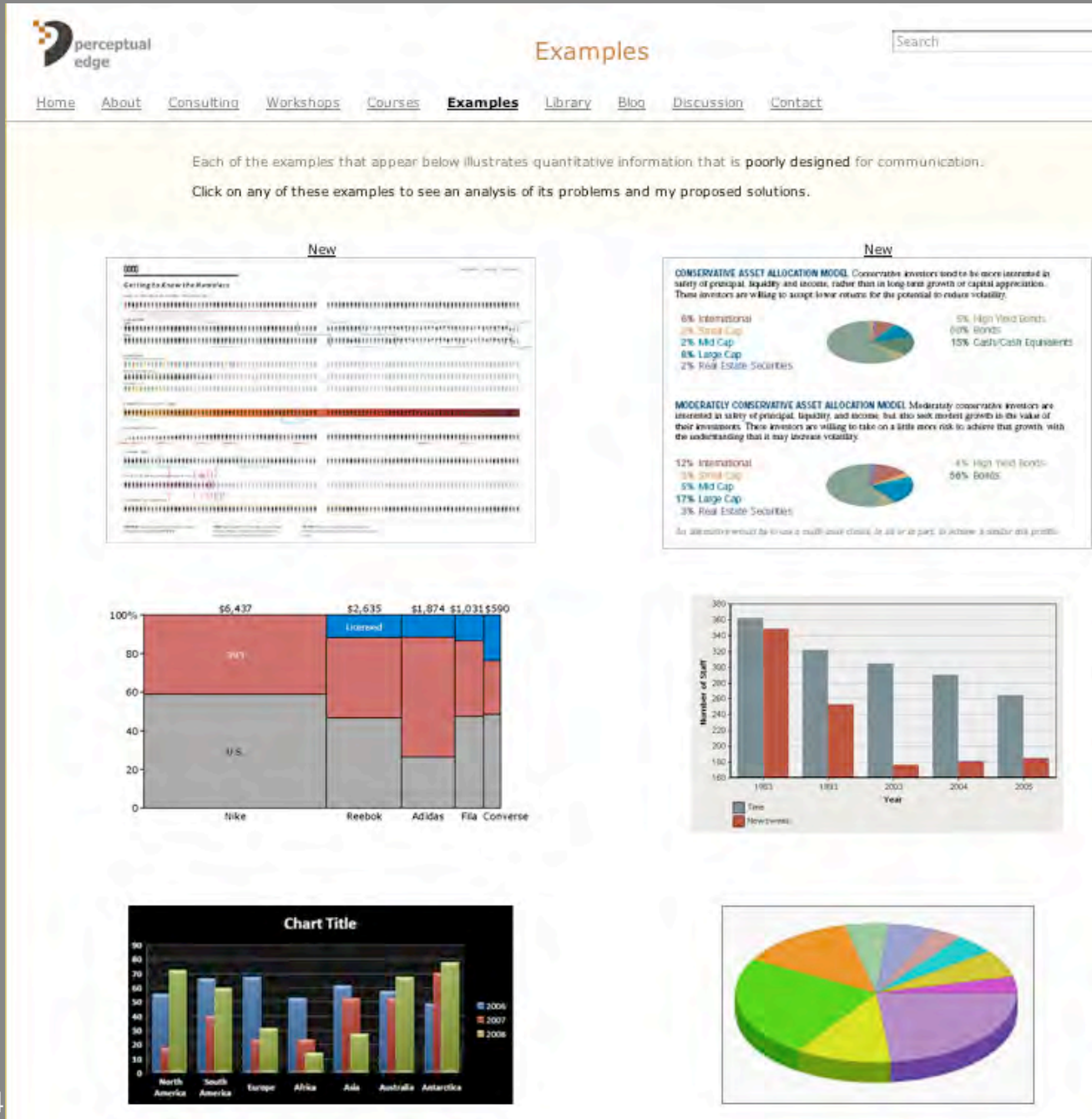
- ✓ Graphical to focus attention on key trends, comparisons and exceptions
- ✓ Display only relevant data
- ✓ Inherently contain predefined conclusions

Note: Collecting user requirements is KEY



*from 'Excel 2007 Dashboards & Reports for Dummies'
by Michael Alexander*

Dashboards - bad examples



Common Problems:

Positioning content in places that don't fit its importance

Positioning content in places that fail to support its use

Including items that serve no useful purpose

Sizing content larger than it deserves

Separating content excessively

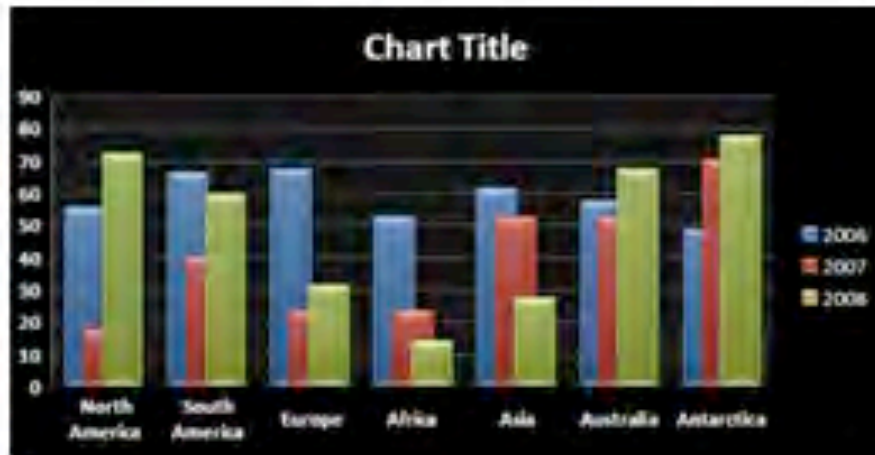
Visually featuring content & other items more than they deserve

Failing to link contents & other items that are related

Visually suggesting links between unrelated content

Enforcing a rigid symmetrical grid

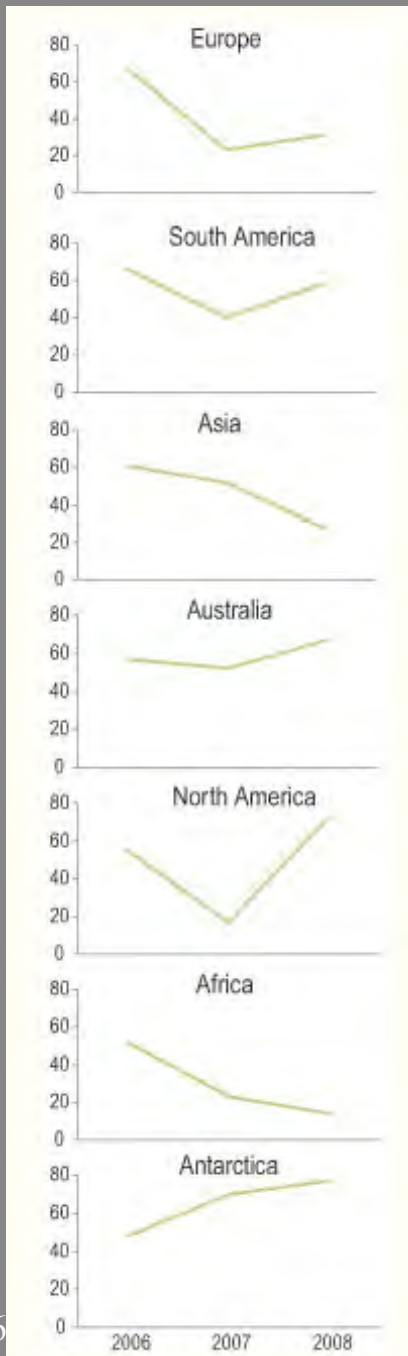
Dashboards - bad examples



Here are a few of this graph's problems:

- There are several distracting (and detracting) visual effects: the reflection of light, transparency, and 3-D effects on the bars (and squares in the legend) add no value.
- The bars have been overlayed on one another, which partially obscures the first two sets and gives them different visual salience. Because the bars for the year 2008 appear in the forefront of each cluster, their greater importance is implied, which was probably not intended. While I can't be sure, the graph's original post date of 2005, suggests that these values are projections, albeit unbelievably volatile ones. Without knowing more about the data, I can't say for sure, but the 2006 projections are probably the surest and most relevant, yet they are partly obscured by the other two years.
- Although the gridlines in this graph are thin and light, because these values are projections, we probably don't need to know precise values. As such, the gridlines are not necessary.
- The bar colors are more intense than they should be. The use of high-intensity colors should be reserved for making important data salient. Regular data should be shown using less intense colors. After all, when you display all of your data to stand out, nothing does.
- The continents have not been ordered in a logical way. At the very least they could have been alphabetized, but, as we'll see below, there's almost always a better way to order your data.
- Although bar graphs are great for showing and comparing the magnitudes of different variables, they are inferior to lines for showing how the values change through time. Because the pattern of change through time is likely more important than the actual magnitudes of the individual values, a line graph would have worked better.

Dashboards - solutions



Line graphs make it especially easy to see the patterns of change and to focus on trends. To avoid the clutter of seven lines on a single graph, I used "small multiples," a series of seven small graphs, which vary by region, but otherwise look and work the same. Small multiples may be arranged vertically (shown above), horizontally, or in a matrix. Because this information is a projection (and so the exact magnitudes are probably not as important), I have made the assumption that the graphs should be arranged to make it easiest to compare the patterns of change for the various regions, which is why I aligned the years by arranging the graphs vertically. If the magnitudes of the lines were more important, then a horizontal layout would have been preferable, for easier magnitude comparisons. Notice that the horizontal label (showing the years) is only shown on the very bottom of the graph. This is all that's necessary to show which part of each line belongs to which year. Duplicating these labels for each graph would have resulted in redundancy and clutter.

I have reordered the continents based on the 2006 values, with the highest at the top and the lowest at the bottom. I based the sequence on the 2006 value because, as these values are projections, the first year is likely to be most reliable and of greatest interest to decision-makers.

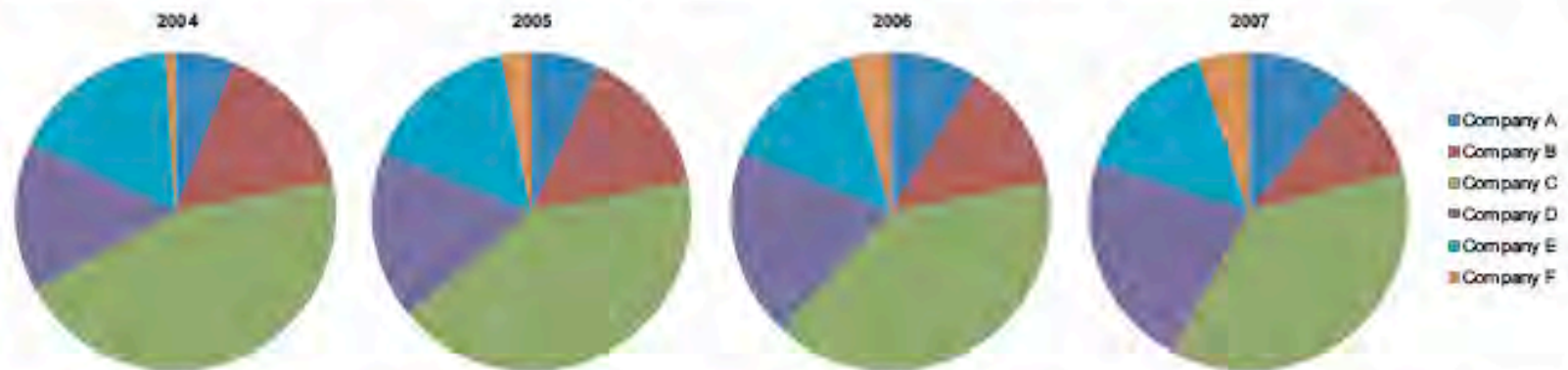
This new design is clean and clear—free of the visual distractions in the first two examples. Anyone viewing the graph would be able to examine the data, focusing perhaps on the large declines that are projected to occur in Europe and Africa, instead of the pretty, shiny bars.

Reduce the non-data ink

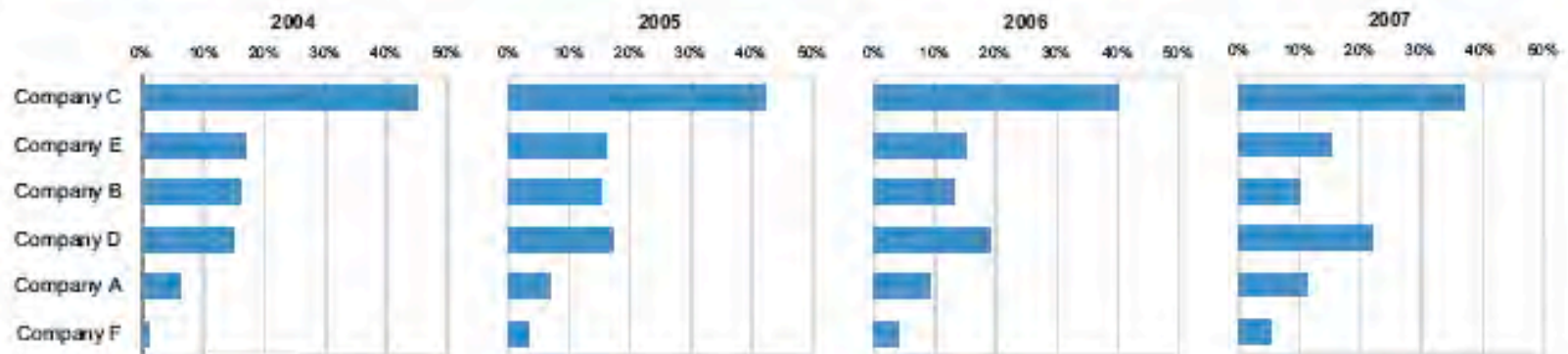
Enhance the data ink

Dashboards - evil pies

Let's examine another ineffective use of pie charts. Edward Tufte once said that "the only worse design than a pie chart is several of them, for then the viewer is asked to compare quantities located in spatial disarray both within and between pies" (Edward Tufte, *The Visual Display of Quantitative Information*, Graphics Press, 1983, p. 178.) I share Tufte's opinion that this is an ineffective way to compare multiple part-to-whole relationships.



Try to follow the changes of these various companies and how they compare to one another through time. It is nearly impossible. Notice how easily you can do it, however, using the following display:



Dashboards - evil pies

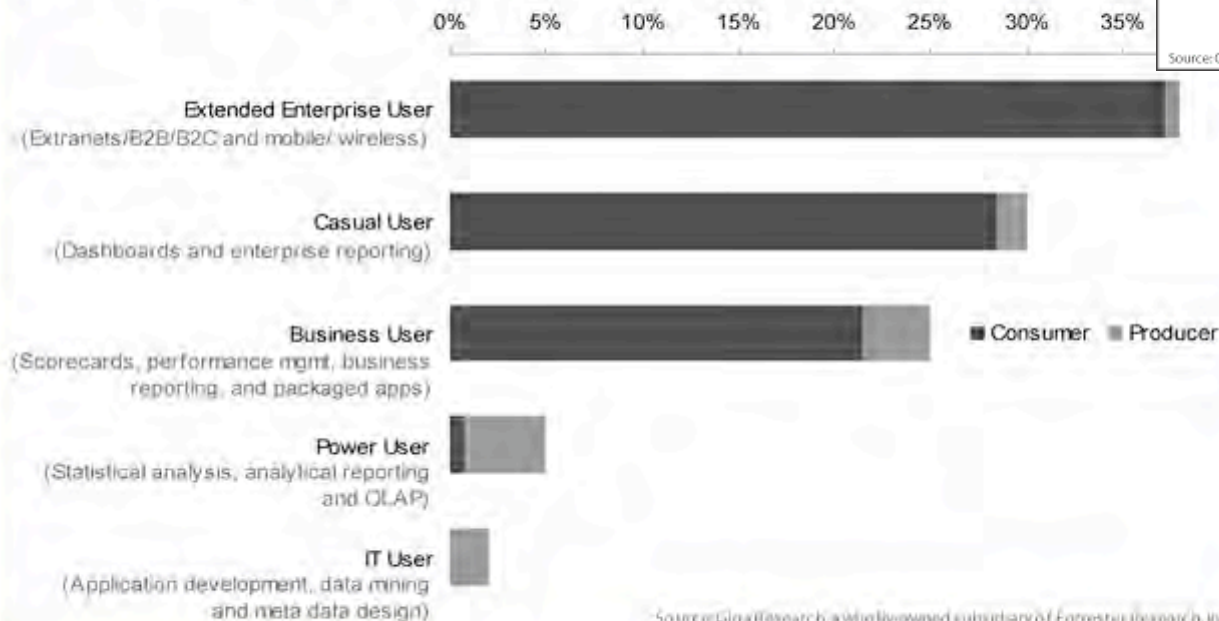
My Analysis

The data is great but the display is a jumbled mess.

A Solution

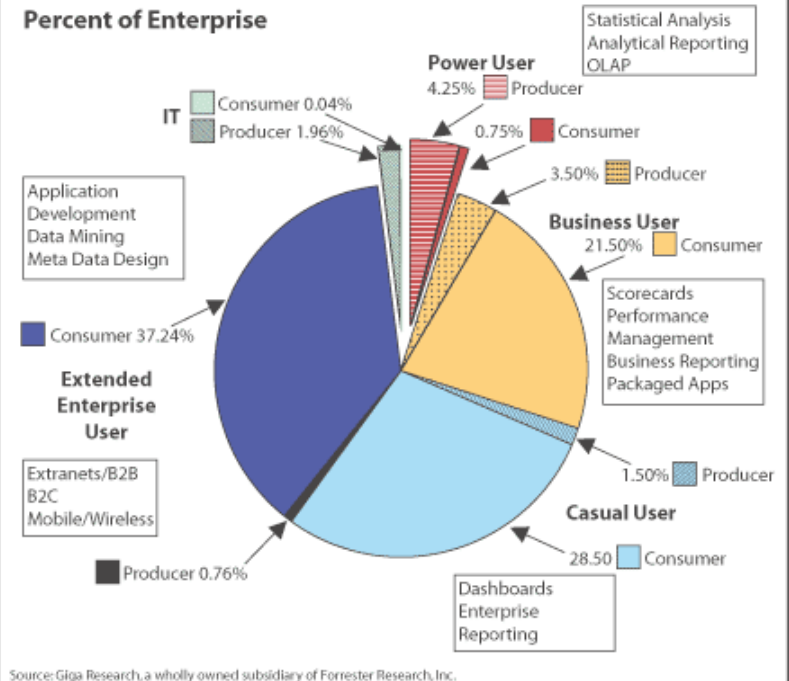
Here's the same data displayed simply and clearly:

Percentage of Analytic Computer Usage by Type



Source: Giga Research, a wholly owned subsidiary of Forrester Research, Inc.

Percent of Enterprise



Source: Giga Research, a wholly owned subsidiary of Forrester Research, Inc.

Reduce the non-data ink
Enhance the data ink

I could have used colors but, frankly, this graph doesn't need them. Limiting it to black and white allows you to photocopy this useful information and pass it on without any loss of quality. Can you imagine what the original pie chart would look like if you photocopied it in black and white?

Dashboards - real estate

“The relative prominence of screen space on a dashboard can be divided into quadrants... Whenever possible, place information that is considered most important in the upper left hand region and that which is least important in the lower right hand corner.”

Emphasized

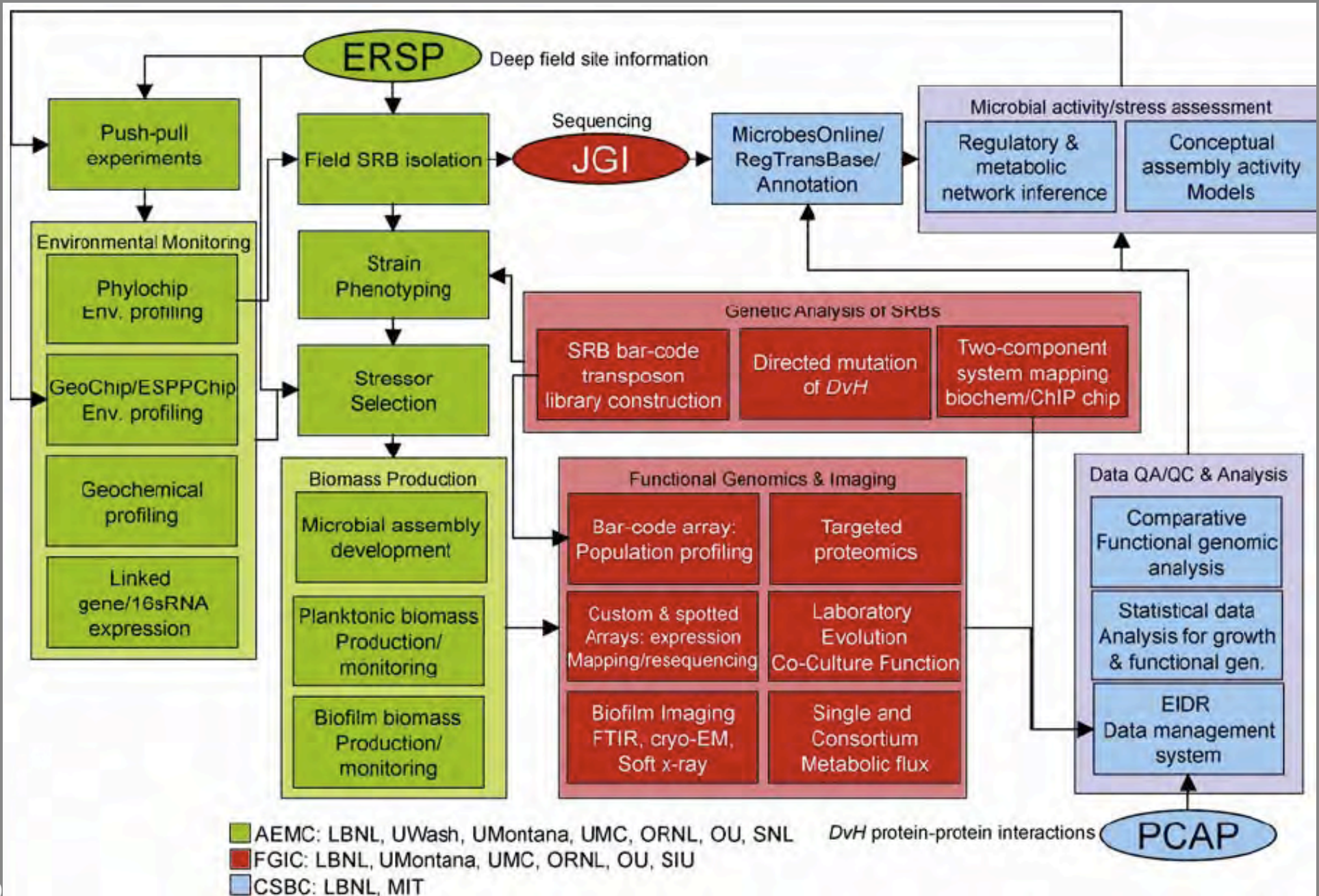
Neither emphasized
nor de-emphasized

Emphasized

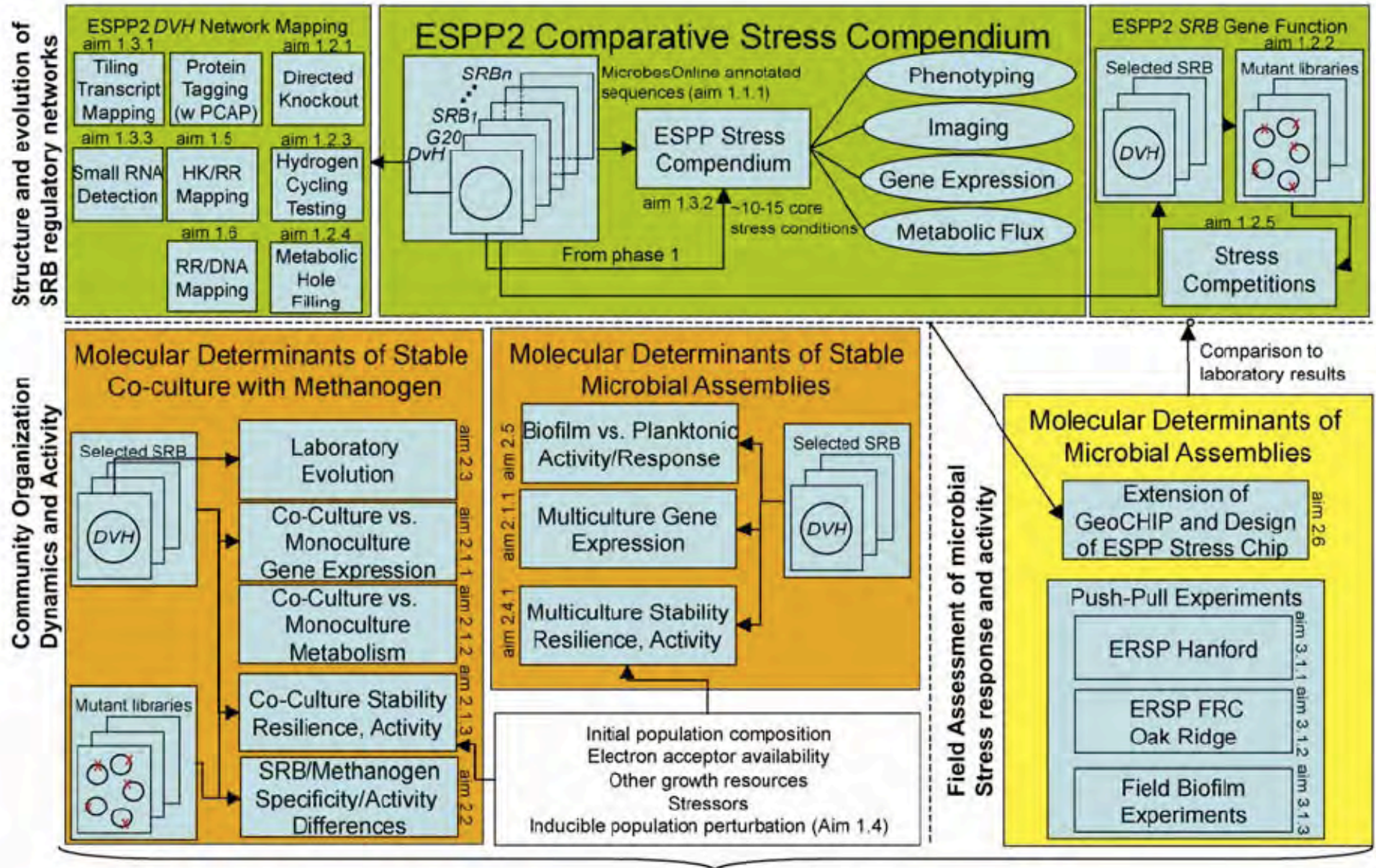
Neither emphasized
nor de-emphasized

De-emphasized

Team Science Approach



Team Science Approach



Milestones: Year One 09/30/2008

AEMC

- Obtain previously isolated SRB (especially for DOE contaminated sites), prepare DNA for sequencing submit to JGI.
- Growth optimization and stability studies of different syntrophic co-culture assemblies: Alternative Dv strains/species.
- Full scale biomass production for steady-state growth stress-perturbed co-culture response experiments (perturbation & steady state analyses using optimized co-culture conditions) for different SRB/methanogen pairs.
- Initial tests of multiculture conditions.
- Initiation of co-culture evolution experiments.
- Optimize transposon strain library competition experiments for read-out by bar code arrays both in monoculture and co-culture.
- Complete membrane profiling of *D. vulgaris* and *M. maripaludis* in mono culture and together in syntrophic culture.
- Design of push-pull experiments and initial characterization of site bacterial populations and geochemistry and Hanford and Oak Ridge, including initial testing of in well sediment/attachment simulation systems.
- Design larger scale attached stress experiments for comparison with planktonic experiments (transcriptomics).
- Complete contrast/compare studies of groundwater and sediment ecogenomics from Oak Ridge site for metagenome (Sanger, 454, and clone libraries), 16SRNA Phylochip, Geochip, and realtime Q-PCR.

FGIC

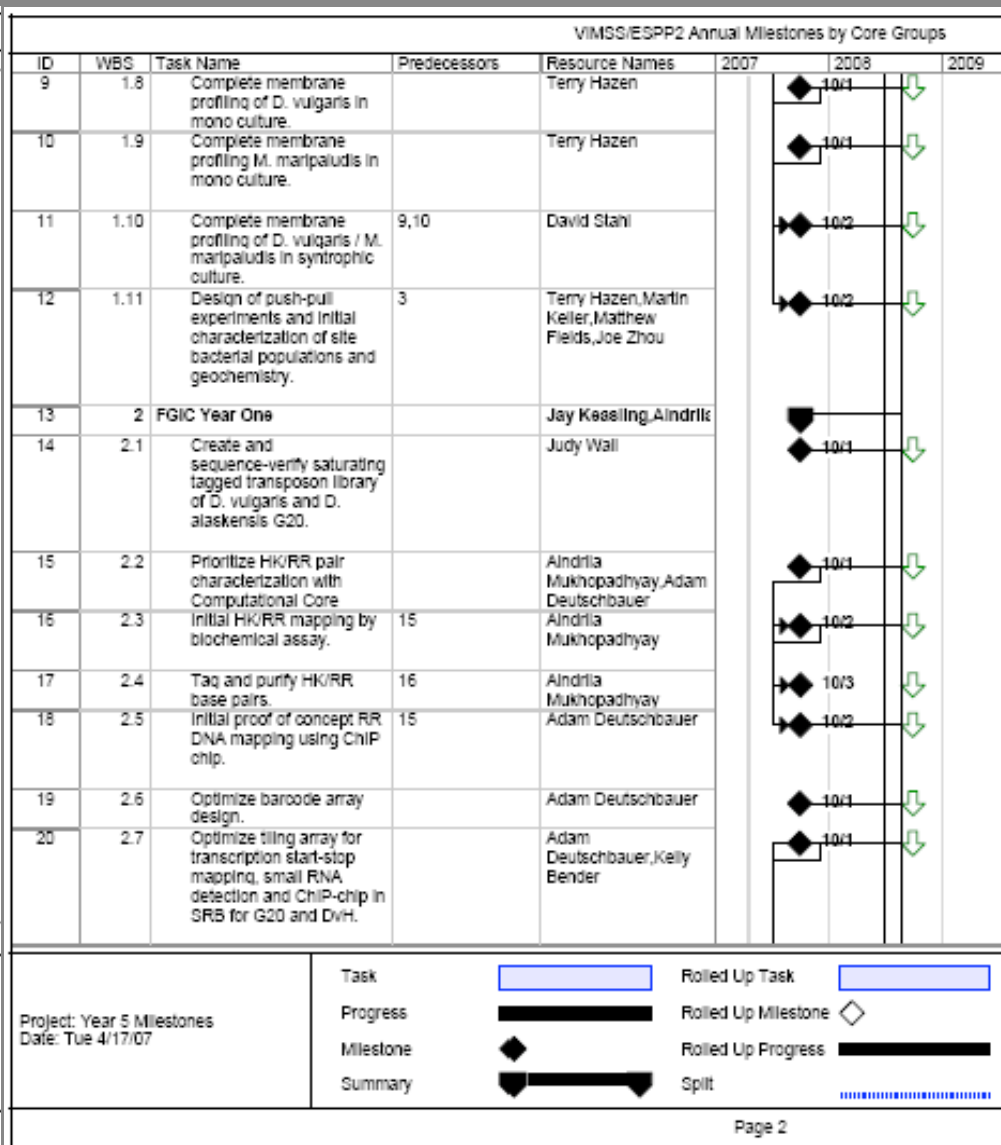
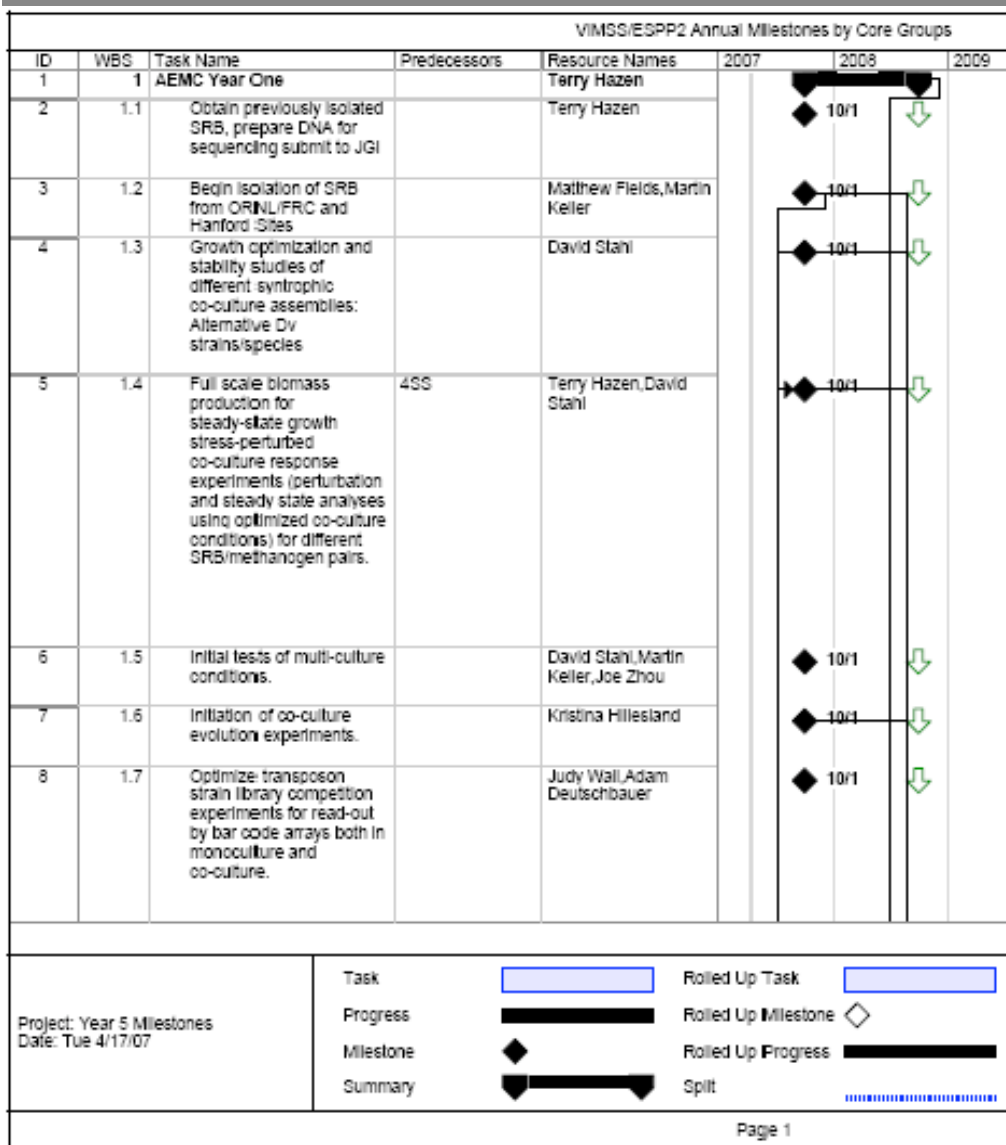
- Create and sequence-verify saturating tagged transposon library of *D. vulgaris* and *D. alaskensis* G20.
- Prioritize HK/RR pair characterization with Computational Core.
- Tag and purify HK/RR pairs.
- Initial HK/RR mapping by biochemical assay.
- Initial proof of concept RR/DNA mapping using ChIP-chip.
- Optimize barcode array design.
- Optimize tiling array for transcription start-stop mapping, small RNA detection and ChIP-chip in SRB for G20 and DvH.
- Optimize multiplex gene expression design for G20.
- Complete stress response transcriptomics for G20.
- Initial survey of possible small RNA regulators.
- Complete design and testing of ESPPChip microarray.

CSBC

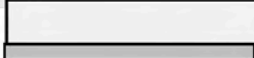


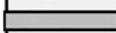



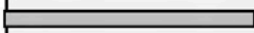


- Extension of MicrobesOnline for 16SRNA, GeoCHIP/ESPPChip, Phenotype, metagenomic data.
- Complete computational analysis of DvH and G20 and methanogen metabolism.
- Establish flux model analysis methods for mono- and multicultures.
- Developing tiling array and bar-code array design and analysis techniques.
- Complete annotation of Dv Miyazaki, Ds 27774, and one Dv Hanford isolate.
- Complete initial reannotation of DvH.
- Begin design of conceptual model of stress, ED, TEA responses for Hanford Cr and Oak Ridge U contaminated sites.

Work Breakdown Structure by Milestones

Schedule Development & Execution



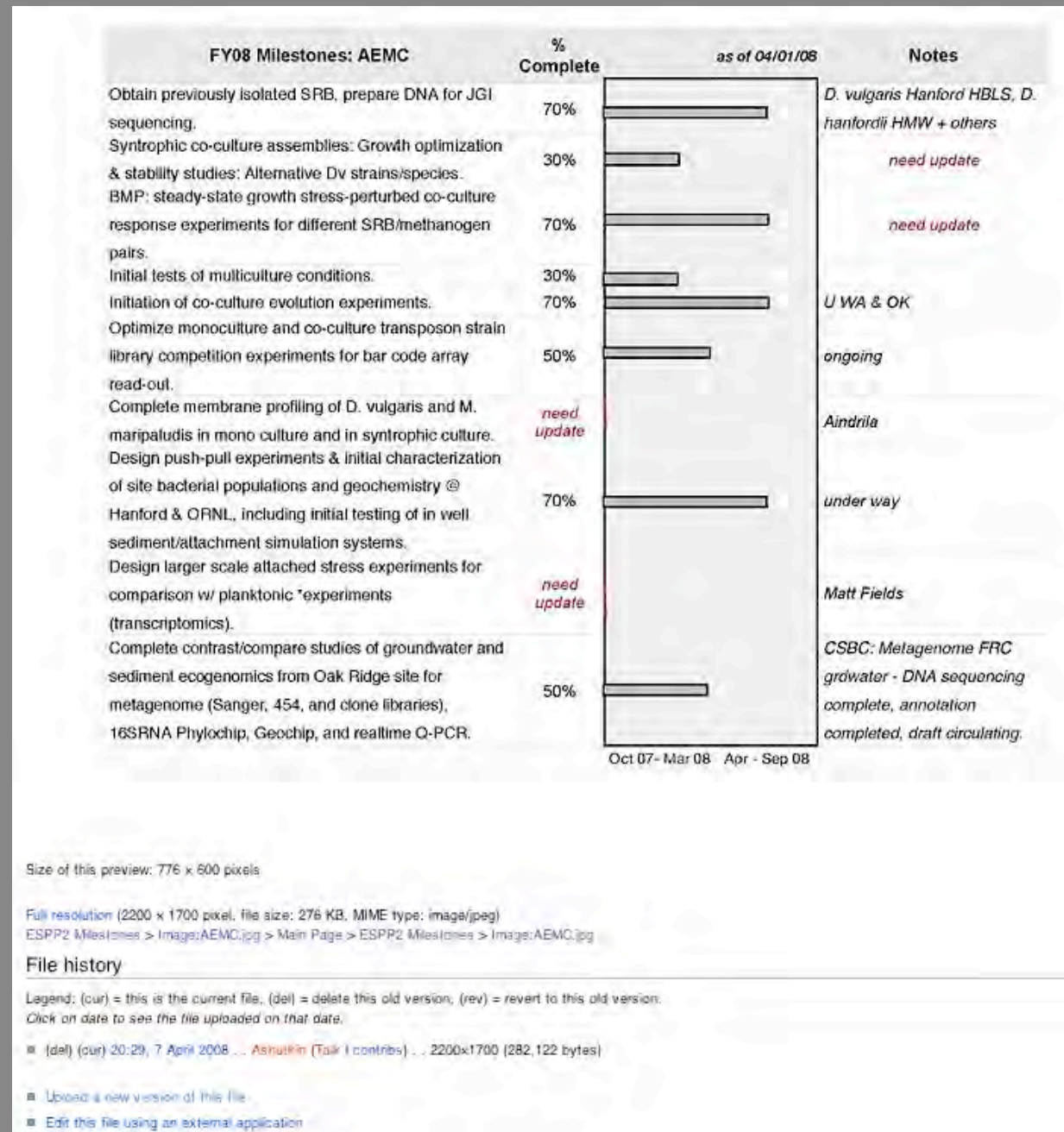
Dashboard Milestone Reports

FY08 Milestones: AEMC	% Complete	as of 04/01/08		Notes
Obtain previously isolated SRB, prepare DNA for JGI sequencing.	70%			<i>D. vulgaris</i> Hanford HBLS, <i>D. hanfordii</i> HMW + others
Syntrophic co-culture assemblies: Growth optimization & stability studies: Alternative Dv strains/species.	30%			<i>need update</i>
BMP: steady-state growth stress-perturbed co-culture response experiments for different SRB/methanogen pairs.	70%			<i>need update</i>
Initial tests of multiculture conditions.	30%			
Initiation of co-culture evolution experiments.	70%			U WA & OK
Optimize monoculture and co-culture transposon strain library competition experiments for bar code array read-out.	50%			ongoing
Complete membrane profiling of <i>D. vulgaris</i> and <i>M. maripaludis</i> in mono culture and in syntrophic culture.	<i>need update</i>			Aindrila
Design push-pull experiments & initial characterization of site bacterial populations and geochemistry @ Hanford & ORNL, including initial testing of in well sediment/attachment simulation systems.	70%			under way
Design larger scale attached stress experiments for comparison w/ planktonic *experiments (transcriptomics).	<i>need update</i>			Matt Fields
Complete contrast/compare studies of groundwater and sediment ecogenomics from Oak Ridge site for metagenome (Sanger, 454, and clone libraries), 16SRNA Phylochip, Geochip, and realtime Q-PCR.	50%			CSBC: Metagenome FRC grdwater - DNA sequencing complete, annotation completed, draft circulating.
		Oct 07- Mar 08	Apr - Sep 08	

Key Performance Indicators (KPI) ~
essential tasks draw attention to problem
areas

ESPP wiki

Dashboard Milestone Reports



Cost Estimation & Budget Oversight

Rate Type FY07 FY08 FY09 FY10 FY11 FY12 (est) 5 Year Totals

Escalation Rates

LBNL Labor 3.5% 3.0% 3.0% 3.0%

LBNL Supplies
& Other
Expenses (OMB)

Total Labor \$3,77

Original Project
Budget
Assumptions

Total Supplies &
Other Expenses
(OMB)

\$4,26

Totals \$8,04

LBNL Labor

Equipment,
Supplies & Other
Expenses, LBNL
(OMB)

LBNL Total

Total Direct Costs,
LBNL

Total Indirect
Costs, LBNL

Total Direct Costs
less Other Inst.
Indirect Costs,
LBNL

SNL

ORNL

ESPP2 Total

% original projection

DOE F-4620-1 (04-23) All Other Editions Are Obsolete		U.S. Department of Energy Budget Page (See reverse for instructions)		
ORGANIZATION Ernest O. Lawrence Berkeley National Laboratory				
PRINCIPAL INVESTIGATOR/PROJECT DIRECTOR Arkin, Adam				
Overall Project Budget				
A. SENIOR PERSONNEL: PPD, Co-PIs, Faculty and Other Senior Associates (List each separately with title, A.S. show number in brackets)		DOE Funded Personnel:		
		CAL	ACAD	SUMR
1. Arkin, Adam		15.00		
2. Keesling, Jay		6.00		
3. Hazen, Jeff		6.00		
4. Duschak, Linda		18.00		
5.				
6. (0) OTHERS (LIST INDIVIDUALLY ON BUDGET EXPLANATION PAGE)		231.00		
7. (11) TOTAL SENIOR PERSONNEL (1-6)		336.00		
B. OTHER PERSONNEL (SHOW NUMBERS IN BRACKETS)				
1. (0) POST DOCTORAL ASSOCIATES		108.00		
2. (25) OTHER PROFESSIONAL (TECHNICIAN, PROGRAMMER, ETC.)		699.00		
3. (2) GRADUATE STUDENTS		54.60		
4. (0) UNDERGRADUATE STUDENTS		0.00		
5. (2) SECRETARIAL - CLERICAL		30.00		
6. (0) OTHER		0.00		
TOTAL SALARIES AND WAGES (A+B)				
C. FRINGE BENEFITS (IF CHARGED AS DIRECT COSTS)				
TOTAL SALARIES, WAGES AND FRINGE BENEFITS (A+B+C)				
D. PERMANENT EQUIPMENT (LIST ITEM AND DOLLAR AMOUNT FOR EACH ITEM)				
TOTAL PERMANENT EQUIPMENT				
E. TRAVEL		1. DOMESTIC (INCL. CANADA AND U.S. POSSESSIONS)		
		2. FOREIGN		
TOTAL TRAVEL				
F. TRAINEE/PARTICIPANT COSTS				
1. STIPENDS (Specify levels, types + totals on budget justification page)				
2. TUITION & FEES				
3. TRAINING TRAVEL				
4. OTHER (Fully explain on justification page)				
TOTAL PARTICIPANTS (0)		TOTAL COST		
G. OTHER DIRECT COSTS				
1. MATERIALS AND SUPPLIES				
2. PUBLICATION COSTS/DOCUMENTATION/COMMUNICATION				
3. CONSULTANT SERVICES				
4. COMPUTER (ADPE) SERVICES				
5. SUBCONTRACTS				
6. OTHER				
TOTAL OTHER DIRECT COSTS				
H. TOTAL DIRECT COSTS (A THROUGH G)				
I. INDIRECT COSTS (SPECIFY RATE AND BASE) Please see the indirect calculation sheets				
TOTAL INDIRECT COSTS				
J. TOTAL DIRECT AND INDIRECT COSTS (H+I)				
K. AMOUNT OF ANY REQUIRED COST SHARING FROM NON-FEDERAL SOURCES				
L. TOTAL COST OF PROJECT (J+K)				

LBNL Forward Pricing Rates Effective October 1, 2006 Revision 5

Rate Type

Escalation Rates

Labor
Supplies & Other Expenses (OMB)
Construction Projects (OECM)

Institutional Rates

General and Administrative
G&A (Off Site) Rate - OFF
Site Support (Fabrication) Rate - FAB
Gretina
Animal Care
General Rate - GR1

LDRD Rate

LDRD Operating and Equipment
(LDRD rate is based on proposed structure, which is currently being reviewed by DOE)

IGPP Rate

IGPP

Procurement Burdens (Base: Cost of procured materials & services)

PO's \$1-\$500,000
PO's Over \$500,000 (per PO)
R&D subcontract's \$1-\$300,000
R&D subcontract's over \$300,000
Intra-University Transfers (IUTs) \$1-\$200,000
Intra-University Transfers (IUTs) over \$200,000
Genomics Procurements (PO's \$1-\$500,000)
Molecular Foundry (PO's \$1-\$500,000)
Project Management (PO's \$1-\$500,000)

Travel (Base: Travel Costs)

Travel Rate

Payroll Burden (Base: Delivered effort cost only)

Career & Term Employees

Post Docs, Visiting Post Docs, Limited Employees, and Visiting Researchers
GSRAs
Students/Rehired Retirees/Employees working variable time
Summer Faculty

Fringe Benefits Only (Base: FTE gross pay only)

Career & Term Employees

Post Docs, Visiting Post Docs, Limited Employees, and Visiting Researchers
GSRAs
Students/Rehired Retirees/Employees working variable time
Summer Faculty

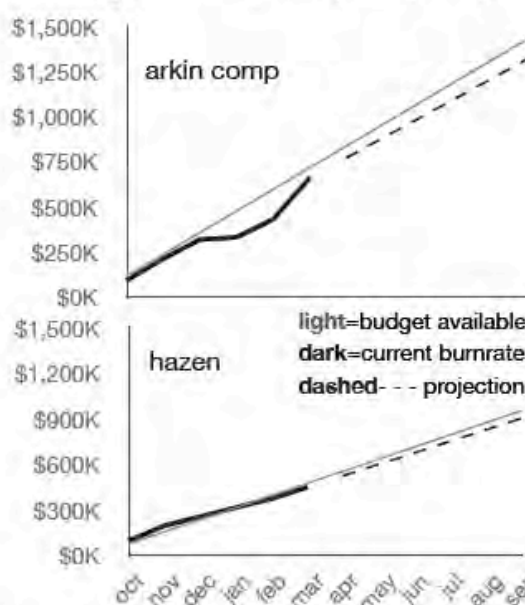
Dashboard Integrated Milestone & Budget Reports

“How then do we make it easy for people to compare related sets of values when they are associated with different units of measure? Two answers come to mind. The first and most obvious is to place them in separate graphs, positioned close to one another so that the patterns in each can be compared to one another, but magnitude comparisons will be discouraged.” *Stephen Few*

light=budget available
dark=current burnrate

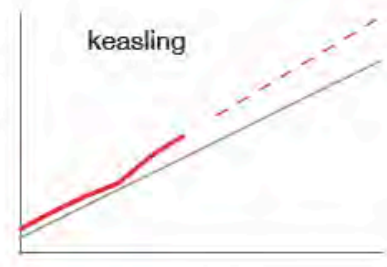


as of 04/01/08

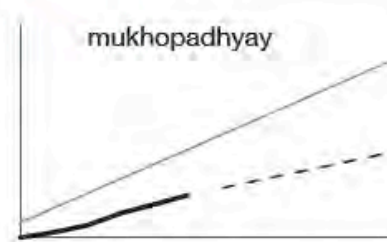


PI	wiki update	Critical Events, Milestones & Top Projects
Arkin - internal for others	3/25/08	Metabolic profile analysis
Hazen	3/25/08	Experimental planning
Keasling	3/31/08	Co-culture Biomass, IMS
Mukhopadhyay	3/12/08	Flux work
Alm	3/24/08	HK knock-outs, HK/RR
Dubchak	2/14/08	Compendium analysis, FRC
Fields	3/25/08	Integrate, annot. metab. pathway
Keller	2/18/08	DvH Cr transcriptomics: biofilm vs planktonic
Singh	3/10/08	Env mRNA profiling, consortia
Stahl	3/7/08	Tri-culture & Mesocosm
	2/29/08	Methanococcus transcriptional/proteomic/mutant analysis in coculture
Wall	1/31/08	Deletions via marker xchange
Bender	2/7/08	Small RNA's - ongoing
Zhou	3/6/08	Evolution, Hanford, FRC
Sequencing	9/7/07	Need update

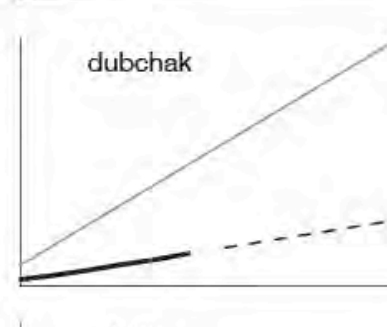
\$580.20K
\$464.16K
\$348.12K
\$232.08K
\$116.04K
\$0K



\$550K
\$440K
\$330K
\$220K
\$110K
\$0K



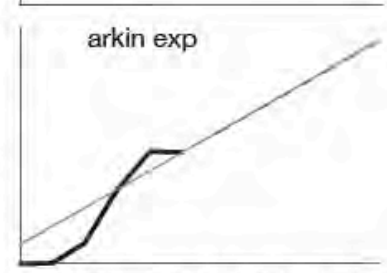
\$300K
\$225K
\$150K
\$75K
\$0K













\$300K
\$225K
\$150K
\$75K
\$0K



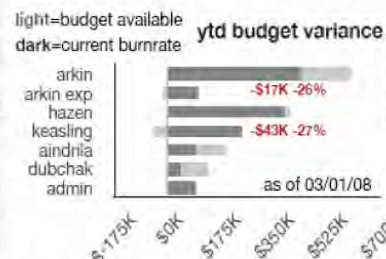
\$225K
\$150K
\$75K
\$0K






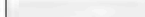










Dashboard Integrated Milestone & Budget Reports

FY08 Milestones: CSBC	% Complete	as of 04/01/08	Notes
Extension of MicrobesOnline for: 16SRNA	50%		
GeoCHIP/ESPPChip	0%		working w/ ORNL to incorporate Selexa data
Phenotype	50%		usefulness will depend on human (undergrads) annotation
Metagenomic data	25%		
Complete computational analysis of DvH & G20 + methanogen metabolism.	50%		Waiting on data.
Establish flux model analysis methods for mono- and multicultures.	50%		High potential value & high interest
Developing tiling array and bar-code array design and analysis techniques.	90%		Analysis techniques will continue to evolve & improve
Complete annotation of Dv Miyazaki, Ds 27774, and one Dv Hanford isolate.	50%		Waiting for these + other sequences before expression array & analysis
Complete initial reannotation of DvH.	50%		Additional data to include? Tiling array?
Begin design of conceptual model of stress, ED, TEA responses for Hanford Cr and Oak Ridge U contaminated sites.	0%		Participating in experimental design - conceptual model depends on data to be collected







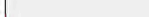
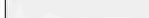

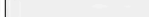
Oct 07 - Mar 08 Apr - Sep 08



PI	wiki update	Critical Events, Milestones & Top Projects
Arkin - internal for others	3/7/08	Microarray analysis
Hazen	3/11/08	Metabolomics condition set
Keasling	3/14/08	Co-culture Biomass, IMS
Mukhopadhyay	3/12/08	Tryptophan biosynthesis
Alm	3/12/08	HK knock-outs, HK/RR
Dubchak	2/14/08	Compendium analysis, FRC
Fields	1/31/08	Hiring Pavel
Keller	2/18/08	DvH Cr transcriptomics: biofilm vs planktonic
Singh	2/7/08	Consortia culture
Stahl	3/7/08	Tri-culture & Mesocosm
Wall	2/29/08	Methanococcus transcriptional/proteomic/ mutant analysis in coculture
Bender	1/31/08	Deletions via marker xchange
Zhou	2/6/08	Small RNA's
Sequencing	3/6/08	Evolution, Hanford, FRC
	9/7/07	Need update

FY08 Milestones: FGIC	% Complete	as of 04/01/08	Notes
Create and sequence-verify saturating tagged transposon libraries for: D. vulgaris	10%		Adam D. just started
D. alaskensis G20	50%		100% created, sequence verification ongoing
Prioritize HK/RR pair characterization w/ CSBC.	30%		
Tag and purify HK/RR pairs.	70%		RR - 90%, HK - 50%
Initial HK/RR mapping by biochemical assay.	50%		Post-doc to begin 08/01. All HK knock-outs: 12/64 genes completed, 2-3 checked. Paper in process.
Initial proof of concept RR/DNA mapping using ChIP-chip.	50%		Ready to go - waiting on Nimblegen chips (Adam D. & Paramvir)
Optimize barcode array design.	50%		(Adam D. & Paramvir)
Optimize tiling array for transcription start-stop mapping, small RNA detection and ChIP-chip in SRB for G20 and DvH.	0%		Will need to recast into year 2 milestones.
Optimize multiplex gene expression design for G20.	need update		Adam D.
Complete stress response transcriptomics for G20.	need update		Adam D.
Initial survey of possible small RNA regulators.	30%		Kelly Bender - 1 study complete
Complete ESPPchip microarray design & testing.	need update		Redundant? Combine w/ ChIP-chip development?
Gene expression compendium	need update		
Metabolomics	need update		

Oct 07 - Mar 08 Apr - Sep 08

FY08 Milestones: AEMC	% Complete	as of 04/01/08	Notes
Obtain previously isolated SRB, prepare DNA for JGI sequencing.	70%		D. vulgaris Hanford HBL5, D. hanfordii HMW + others
Syntrophic co-culture assemblies: Growth optimization & stability studies: Alternative Dv strains/species.	30%		need update
BMP: steady-state growth stress-perturbed co-culture response experiments for different SRB/methanogen pairs.	70%		need update
Initial tests of multiculture conditions.	30%		
Initiation of co-culture evolution experiments.	70%		U WA & OK
Optimize monoculture and co-culture transposon strain library competition experiments for bar code array read-out.	50%		ongoing
Complete membrane profiling of D. vulgaris and M. maripaludis in mono culture and in syntrophic culture.	need update		Aindria
Design push-pull experiments & initial characterization of site bacterial populations and geochemistry @ Hanford & ORNL, including initial testing of in well sediment/attachment simulation systems.	70%		under way
Design larger scale attached stress experiments for comparison w/ planktonic *experiments (transcriptomics).	need update		Matt Fields
Complete contrast/compare studies of groundwater and sediment ecogenomics from Oak Ridge site for metagenome (Sanger, 454, and clone libraries), 16SRNA Phylochip, Geochip, and realtime Q-PCR.	50%		CSBC: Metagenome FRC grdwat - DNA sequencing complete, annotation completed, draft circulating.

Oct 07 - Mar 08 Apr - Sep 08

Adam P. Arkin and Terry C. Hazen, Directors

Applied Environmental Microbiology Core:

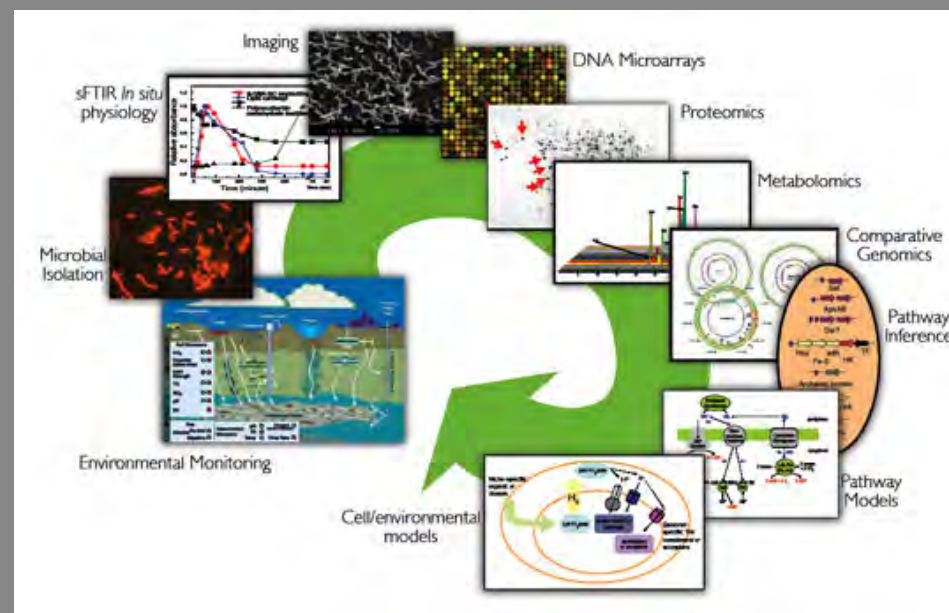
- LBNL, Terry C. Hazen
- University of Washington, David Stahl
- Montana State University, Matthew Fields

Functional Genomics and Imaging Core:

- LBNL, Jay Keasling and Aindrila Mukhopadhyay
- University of Missouri-Columbia, Judy Wall
- Southern Illinois University, Kelly Bender
- Sandia National Laboratory, Anup Singh
- Oak Ridge National Laboratory, Martin Keller
- University of Oklahoma, Jizhong (Joe) Zhou

Computational and Systems Biology Core:

- LBNL: Adam P. Arkin, Inna Dubchak, Paramvir Dehal
- MIT: Eric Alm



ACKNOWLEDGEMENT

ESPP2 is part of the Virtual Institute for Microbial Stress and Survival supported by the U. S. Department of Energy, Office of Science, Office of Biological and Environmental Research, Genomics Program:GTL through contract DE-AC02-05CH11231 between Lawrence Berkeley National Laboratory and the U. S. Department of Energy.